

UNITED STATES PATENT APPLICATION  
FOR  
A METHOD AND SYSTEM FOR ENFORCING CONSISTENT  
PER-PHYSICAL PAGE CACHEABILITY ATTRIBUTES

INVENTORS:  
ALEXANDER KLAIBER  
DAVID DUNN

Prepared by:  
WAGNER, MURABITO & HAO, L.L.P.

TWO NORTH MARKET STREET  
THIRD FLOOR  
SAN JOSE, CALIFORNIA 95113  
(408) 938-9060

# A METHOD AND SYSTEM FOR ENFORCING CONSISTENT PER-PHYSICAL PAGE CACHEABILITY ATTRIBUTES

## TECHNICAL FIELD

5           Embodiments of the present invention relate to the field of memory management. Specifically, embodiments of the present invention relate to a method and system for enforcing consistent per-physical page cacheability attributes.

## 10   BACKGROUND OF THE INVENTION

          Embodiments of the present invention relate to the field of memory management. Specifically, embodiments of the present invention relate to a method and system for enforcing consistent per-physical page cacheability attributes.

15

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A depicts an exemplary dual path processor access to a memory, according to one embodiment of the present invention.

5           Figure 1B depicts an exemplary alias collision in a TLB with the dual path processor access to a memory.

            Figure 1C depicts an exemplary snooping mechanism to maintain cache coherency in a dual path processor access to a memory, according to one  
10   embodiment of the present invention.

            Figure 2 depicts an exemplary table for indexing physical memory pages and indicating their actual cacheability attribute, according to one embodiment of the  
15   present invention.

            Figure 3 depicts an exemplary candidate TLB entry with a desired cacheability state component, according to one embodiment of the present  
invention.

20           Figure 4 depicts a system for comparing an actual cacheability characteristic and a desired cacheability state, according to one embodiment of the present  
invention.

            Figure 5 depicts a system for enforcing consistent cacheability characteristics,  
25   according to one embodiment of the present invention.

Figure 6 is a flowchart of a process for comparing the desired cacheability attribute of a page to an actual cacheability status attributed to that page address, according to one embodiment of the present invention.

5            Figure 7 is a flowchart of a process for accessing desired cacheability states and current cacheability status and comparing them, according to one embodiment of the present invention.

            Figure 8 is a flowchart of a process for determining a cacheability transition  
10    direction, according to one embodiment of the present invention.

            Figure 9 is a flowchart of a process for handling a cacheability exception in a cacheable-to-non-cacheable transition, according to one embodiment of the present invention.  
15

            Figure 10 is a flowchart of a process for handling a cacheability exception in a non-cacheable-to-cacheable transition, according to one embodiment of the present invention.

## DETAILED DESCRIPTION OF THE INVENTION

A method and system for enforcing consistent per-physical page cacheability attributes is disclosed. Reference will now be made in detail to embodiments of the invention, examples of which are illustrated in the accompanying drawings.

5 While the invention will be described in conjunction with these embodiments, it will be understood that they are not intended to limit the invention to these embodiments. On the contrary, the invention is intended to cover alternatives, modifications, and equivalents, which may be included within the spirit and scope of the invention as defined by the appended claims.

10

Furthermore, in the following detailed description of the present invention, numerous specific details are set forth in order to provide a thorough understanding of the present invention. In other instances, well known components, circuits, methods, materials, and procedures have not been described in detail so as not to unnecessarily obscure aspects of the present invention. Embodiments of the  
15 present invention are discussed primarily in the context of a method and system for enforcing consistent per-physical page cacheability attributes.

Modern computer system memory systems typically utilize multiple storage  
20 levels, including caching and paging. In paging, physical memory is divided into page frames of a fixed size. A page table (typically residing in main memory), comprising one or more Page Table Entries (PTEs), translates virtual addresses into physical addresses used to access memory. To make the translation process more efficient and thus reduce latency, a translation lookaside buffer (TLB) can be used to  
25 cache a subset of the page table.

Some processors access memory via multiple paths. 'X86' processors for instance can employ two separate paths. Figure 1A depicts an exemplary processor 10. Core 14 can access memory 21 either via cacheable path 11 or non-cacheable path 12. Cacheable path 11 has a first cache L1 and a second cache L2.

5 Non-cacheable path 12 may include write combiners 16.

Having dual paths between core 14 and memory 21 can introduce cache coherency issues, especially as multiple virtual addresses can map to the same physical address, as depicted in Figure 1B. Within TLB 30 and PTE 39 of page

10 table 40, an entry corresponding to cacheable path 11 associates the virtual address (VAddr) 31 with physical address (PAddr) 34. However, another PTE corresponding to non-cacheable path 12 associates the VAddr 32 with PAddr 34 in PTE 39 and TLB 30.

15 In the case described above, when the same physical page is accessible via both a cacheable path 11 and a non-cacheable path 12, additional action should be taken to preserve cache coherency and/or memory ordering. In prior art, cache coherency can be maintained by snooping. As depicted in Figure 1C, when core

20 14 is accessing via non-cacheable path 12, snoop 49 examines the caches L1 and L2 on a per cache line basis to determine whether a copy is present therein. If a copy is detected in the caches L1 or L2, the data is written back before performing the non-cacheable access, such that the non-cacheable access observes the latest copy of the data from cacheable path 11. Conversely, any pending writes in non-cacheable path 12 (e.g., in write combiners 16) are drained (i.e., written back to main

25 memory) before accessing data through cacheable path 11.

Snoop 49 intertwines cacheable path 11 and non-cacheable path 12.

Intertwining can impact hardware resources, because dedicated hardware resources for performing snoops is required or hardware resources occupied by detailing to snoop are unavailable to perform other tasks while snooping. Further, intertwining  
5 the cacheable and non-cacheable paths makes it difficult to independently optimize them for speed.

A processor may operate in an *aggressive* execution mode wherein it performs *speculative* memory updates, or a *conservative* execution mode wherein  
10 the processor does not perform speculative memory updates. Speculative updates may be *committed* at certain points in the execution, or undone by a *rollback* operation. In one embodiment, speculative memory updates may be stored in the L1 cache. In such processors, it may not in fact be permissible to flush speculative data from the caches. In such a case, snoop 49 is deterred from  
15 operating.

Certain portions of the detailed descriptions of embodiments of the invention, which follow, are presented in terms of processes and methods (e.g., processes 600, 700, 800, 900, and 1000 of Figure 6, 7, 8, 9, and 10, respectively). Although  
20 specific steps are disclosed herein describing the operations of these processes and methods, such steps are exemplary. That is, embodiments of the present invention are well suited to performing various other steps or variations of the steps and/or order of steps recited in the flowcharts of the figures herein.

25 In one embodiment, such processes are carried out by processors and electrical/electronic components under the control of computer readable and computer executable instructions comprising code contained in a computer usable

medium. The computer readable and computer executable instructions reside, for example, in code within a computer usable medium and used in the processor, data storage features, memory, registers and other components of a computer system performing the method for maintaining sufficient bandwidth on a network link during  
5 failures. However, the computer readable and computer executable instructions may reside in any type of computer readable medium.

A computer system that embodies a system and performs a method for enforcing consistent per-physical page cacheability attributes can comprise any kind  
10 of computer system with sufficient computing power and memory capacity. For example, the computer system can comprise a workstation computer system, an embedded system, a personal computer system, a specialized computing system, a main-frame computer system, or a supercomputer system. Modules of the system for enforcing consistent per-physical page cacheability attributes can be  
15 implemented in software, firmware, and/or hardware or any combination of software, firmware, and/or hardware.

A method and system for enforcing consistent per-physical page cacheability attributes are disclosed. In one embodiment, the method for enforcing consistent  
20 per-physical page cacheability attributes maintains memory coherency within a processor addressing memory by comparing a desired cacheability attribute of a physical page address in a page translation entry (PTE) against an authoritative table (e.g., physical map) that indicates the current cacheability status. This comparison is made, in one embodiment, at the time the PTE is inserted into a translation lookaside  
25 buffer (TLB). In one embodiment, when the comparison detects a mismatch between the desired cacheability attribute of the page and the page's current cacheability status, corrective action can be taken to transition the page into the



desired cacheability state. In one embodiment, the corrective action can be performed by either software or hardware.

Therefore, hardware-intensive snooping methods can be dispensed with.

5 The method for enforcing consistent per-physical page cacheability attributes maintains memory coherency within a processor addressing memory without costliness in terms of its impact on existing hardware resources. The method for enforcing consistent per-physical page cacheability attributes also maintains memory coherency within a processor addressing memory that does not require flushing a  
10 cache that contains speculative cache lines or waiting for a commit. Further, the method for enforcing consistent per-physical page cacheability attributes implements a cache coherence model within a processor addressing memory that does not rely on snooping and that does not intertwine cacheable and non-cacheable paths.

15

#### EXEMPLARY CACHEABILITY TABLE

Figure 2 depicts an exemplary table for indexing physical memory pages and indicating their actual cacheability attribute, according to one embodiment of the  
20 present invention. Physical map 202 maps memory pages of physical memory 201 according to their identity and identifies the pages' respective current cacheability characteristic. The current cacheability characteristic of a page in physical memory 201 corresponds to the route taken to access that page within the physical memory.

25

A page accessed via a cacheable path (e.g., cacheable path 11; Figure 1) has a current cacheability characteristic of "cacheable." Having a current cacheability

characteristic of cacheable reflects the fact that a memory page is eligible to be cached within a processor. Having a current cacheability characteristic of non-cacheable reflects the fact that a memory page is not cached within the processor and is accessed via a non-cacheable path (e.g., non-cacheable path 12; Figure 1).

5

Physical map 202 can comprise a table in a private memory such as a code morphing software memory. Each page within physical memory 201 is indexed by its identifier, which can comprise a page number, within physical map 202.

Exemplary page 204 is indexed within physical map by a page number entry 205  
10 corresponding to its page number. Within physical map 202, page number entry 205 is associated with a cacheability bit 203.

Associated with page number entry 205, cacheability bit 203 identifies the current cacheability characteristic of the memory page 204. Cacheability bit 203  
15 indicates that the page was last used in either a cacheable or a non-cacheable way. Thus, cacheability bit 203 states whether the corresponding page in physical memory is currently cacheable or non-cacheable.

#### EXEMPLARY TLB ENTRY REFLECTIVE OF DESIRED CACHEABILITY

20 Figure 3 depicts an exemplary translation lookaside buffer (TLB) entry 300, according to one embodiment of the present invention. TLB entry 300 maps a virtual address 305 to a physical address 301 of a page of physical memory (e.g., page 204; Figure 2), with any associated permission bits 302. Further, the TLB entry 300 comprises desired cacheability state 303 corresponding to the physical  
25 memory page associated with virtual address 305.

Desired cacheability state 303 represents the processor desired path for accessing physical memory. For instance, where a processor will be using data being written during the write, a cacheable path (e.g., cacheable path 11; Figure 1) is desirable. Thus, the desired cacheability state corresponding to the memory page being written is "cacheable." By contrast, where a processor will not be using data being written during the write, such as when writing graphics data to a GPU a non-cacheable path (e.g., cacheable path 12; Figure 1) is desirable. Thus, the desired cacheability state corresponding to the memory page being written is "non-cacheable."

10

#### EXEMPLARY CONSISTENT CACHEABILITY CHARACTERISTICS ENFORCER

Figure 4 depicts an exemplary system 400 for enforcing consistent cacheability characteristics, according to one embodiment of the present invention. Consistent cacheability characteristics enforcer 400 comprises a cacheability comparator 401, which operates in one embodiment at TLB insertion time. Before inserting a candidate entry into TLB 410, system 400 obtains cacheability bit 203 from physical map 202. Cacheability comparator 401 compares cacheability bit 203 with desired cacheability state 303, read from the candidate entry to TLB 410. The physical map 202 is indexed using an identifier such as the physical page number of the page as obtained from the TLB 410.

15

20

Upon comparing the current cacheability characteristic reflected by cacheability bit 203 with the desired cacheability state 303 of the TLB 410 entry, three results are possible in one embodiment. The desired cacheability state 303 can either match the current cacheability characteristic reflected by cacheability bit 203, i.e., both cacheable or both non-cacheable, or one of two possible mismatches is detected.

25

The first cacheability mismatch indicates that the current cacheability characteristic reflected by cacheability bit 203 is "cacheable," but that the desired cacheability state 303 is "non-cacheable." This mismatch is indicative that a cacheable-to-non-cacheable transition is to occur in accessing the data associated with the new TLB entry. To enforce memory coherency upon this transition, in one embodiment currently cached copies of the physical page are flushed back to main memory. This flushing can be done upon demand. The candidate entry is then added to the TLB 410 and any old entries that map to the affected physical address are discarded, and the physical map 202 is updated to reflect the non-cacheable mode.

The second cacheability mismatch indicates that the current cacheability characteristic reflected by cacheability bit 203 is "non-cacheable," but the desired cacheability state 303 is "cacheable." This mismatch is indicative that a non-cacheable-to-cacheable transition is to occur in writing the data associated with the new TLB entry, so as to match the processor's cacheability model. In one embodiment, no action is taken. However, in this embodiment, stale data can occur, as outstanding writes may still exist in the non-cacheable path (e.g., write combiners 16; Figure 1). In another embodiment, the outstanding writes are drained to main memory before continuing. The candidate entry is added to TLB 410 and any old entries that map to the affected physical address are discarded, and the physical map 202 is updated to reflect the cacheable mode.

#### Exemplary Cacheable to Non-Cacheable Transition

In one embodiment, upon ascertaining that a cacheable-to-non-cacheable transition is to occur, an exception occurs as a result of the mismatch and then it is determined whether the current execution mode is conservative or aggressive. If

the execution mode is aggressive, system 400 rolls back to the last commit point and enters conservative mode. The rollback effectively discards all speculative data accumulated to that point. After entry into conservative execution mode, execution is retried, causing System 400 to run forward without speculative updates to  
5 memory until the same exception point is reached. At that point, system 400 handles the execution according to the below fix-up operation.

In conservative execution mode, consistent cacheability characteristics enforcer 400 in one embodiment performs a cacheability fix-up operation. The fix-  
10 up operation can comprise removing any old entries in TLB 410 to the physical address currently to be accessed, updating physical map 202 such that the cacheability bit 303 reflects "non-cacheable," and flushing the cache (which, on account of executing in conservative mode, contains no speculative data) of the affected physical page. Execution then continues with entering the candidate entry  
15 into TLB 410 and eventually resuming the instruction that caused the miss in TLB 410.

#### Exemplary Cacheable to Non-Cacheable Transition

In one embodiment, upon ascertaining that a non-cacheable-to-cacheable  
20 transition is to occur, an exception is called as a result of the mismatch and then it is determined whether the execution mode is conservative or aggressive. If the execution mode is aggressive, system 400 rolls back to the last commit point and enters conservative mode. System 400 then runs forward in conservative mode until the same exception is reached. At that point, system 400 handles the  
25 execution according to the below fix-up operation.

In conservative execution mode, consistent cacheability characteristics enforcer 400 in one embodiment performs a cacheability fix-up operation. The fix-up operation can comprise draining to memory all pending writes in the non-cacheable path, for example from write combiners therein (e.g., non-cacheable path 12, write combiners 16; Figure 1A), removing any old entries in TLB 410 to the physical address currently being accessed, and updating physical map 202 such that the cacheability bit 303 reflects "cacheable." Execution then continues with entering the candidate entry into TLB 410 and eventually resuming the instruction that caused the miss in TLB 410.

10

#### EXEMPLARY SYSTEM

Figure 5 depicts a system for enforcing consistent cacheability characteristics, according to one embodiment of the present invention. In one embodiment, the consistent cacheability characteristics enforcer comprises a system 400 that interacts with memory 201 and its physical map 202 to allow the exemplary transitions described above. Figure 5 depicts system 400 deployed within processor 499. At the time an entry is made into TLB 410 (i.e., TLB insertion time), a physical address within memory 201 corresponding to a virtual address supplied by processor 499 is obtained from page table entry (PTE) 519 within a page table 520. A walker-extractor 545 walks page table 520 and extracts therefrom a PTE 519 corresponding to the physical address.

Walker-extractor 545 provides PTE 519 to TLB 410. TLB 410 provides the desired cacheability state 302 derived or extracted from PTE 519 to cacheability comparator 401. Cacheability comparator 401 ascertains the current attributed cacheability status 303a by accessing cacheability bit 303b associated with the physical address within physical map 202. Where the desired cacheability 302

matches the current attributed cacheability 303a, a cacheability match 512 is reported to TLB miss handler 535. TLB miss handler 535 can then proceed to handle the miss for TLB 410 according to techniques known in the art.

5           Where it determines that the desired cacheability 302 does not match the actual cacheability 303a, cacheability comparator 401 can generate an exception 513. Exception 513 is handled by exception handler 520. Exception handler 520 can, in one embodiment, take no action. In one embodiment, exception handler 520 can operate within processor 499 to handle the exception by causing a cacheable-  
10 to-non-cacheable transition or a non-cacheable-to-cacheable transition, as discussed above for instance.

#### EXEMPLARY COMPUTER IMPLEMENTED PROCESSES

Figure 6 is a flowchart of an exemplary computer implemented process for  
15 comparing the desired cacheability attribute of a page to an actual cacheability status attributed to that page address at TLB insertion time, according to one embodiment of the present invention. Process 600 can be invoked in response to a candidate TLB entry to be added to the TLB. Process 600 begins with step 601, wherein an identifier of the page of physical memory, such as a page number, and the desired  
20 cacheability state thereof is extracted or derived from the candidate TLB entry.

In step 602, the current cacheability characteristic of the physical memory page (of the candidate TLB entry) is retrieved by mapping the physical memory page according to an identifier thereof (e.g., a page number) to an entry in an  
25 authoritative table (e.g., physical map 202; Fig. 2). The table lists the page by its identifier along with its current attributed cacheability. The table can be deployed within a private memory.

In step 603, the desired cacheability state and the current cacheability characteristic are compared. In step 604, it is determined whether the desired cacheability state and the current cacheability characteristic match. If the desired  
5 cacheability state and the current cacheability characteristic match, then in step 605, the candidate TLB entry is added to the TLB. At this point, TLB exception handling can be performed by techniques known in the art. If the desired cacheability state and the current cacheability characteristic do not match, then in step 606 in one embodiment, an exception is taken. In step 607, cacheability coherency is fixed-up,  
10 such as by operations described below (e.g., Processes 800, 900, 1000; Figures 8, 9, 10, respectively). Upon fixing-up the cacheability coherency, step 605 is executed, wherein the candidate TLB entry is added to the TLB. Process 600 can be complete upon completion of step 605.

15 Figure 7 is a flowchart of an exemplary computer implemented process for accessing desired cacheability states and current cacheability status and comparing them, according to one embodiment of the present invention. Process 700 can be invoked in response to a candidate TLB entry to be added to the TLB. Process 700 begins with step 701 for walking a page table within a physical memory using a  
20 virtual address to find a PTE corresponding to a page of the physical memory.

In step 702, the appropriate PTE is extracted and read into a TLB. In step 703, the TLB determines an identifier such as a page number and the desired cacheability state for the page of physical memory represented by that PTE. In  
25 step 704, a cacheability map of the pages comprising a physical memory is indexed using the identifier for the page of physical memory. In step 705, the



current cacheability characteristic attributed to that page and associated within the cacheability map with that identifier is determined.

In step 706, the desired cacheability state and the current cacheability  
5 characteristic are compared. In step 707, it is determined whether the desired cacheability state and the current cacheability characteristic match. If the desired cacheability state and the current cacheability characteristic match, then in step 708, the candidate entry is added to the TLB. Normal TLB miss handling can then performed by techniques known in the art. If the desired cacheability state and the  
10 current cacheability characteristic do not match, then in step 709, an exception is generated. Corrective action is taken in one embodiment by an operation as described below (e.g., Process 800; Figure 8). Process 700 can be complete upon completion of steps 708 or 709.

15 Figure 8 is a flowchart of an exemplary computer implemented process 800 for determining a cacheability transition direction, according to one embodiment of the present invention. Process 800 begins with step 801, wherein a mismatch is detected between the desired cacheability state of a candidate for entry into a TLB and the current cacheability characteristic attributed to the corresponding physical  
20 memory page. The current cacheability characteristic can be obtained for instance from an authoritative table (e.g., physical map 202; Figure 2), which maps each page of a physical memory and indexes cacheability characteristics with each page map entry.

25 In step 802, a desired cacheability transition is determined. For instance, if the mismatch detected is between a desired cacheability state of non-cacheable and a current attributed cacheability characteristic of cacheable, then the desired cacheability

transition is from cacheable to non-cacheable. If on the other hand the mismatch detected is between a desired cacheability state of cacheable and a current attributed cacheability characteristic of non-cacheable, then the desired cacheability transition is from non-cacheable to cacheable.

5

Upon determining that the desired cacheability transition is from cacheable to non-cacheable, then in step 803, it is determined whether the execution mode is conservative or aggressive. Upon determining that the desired cacheability transition is from non-cacheable to cacheable, then in step 804, it is determined  
10 whether the execution mode the execution mode is conservative or aggressive.

If it is determined that the execution mode is aggressive, then in step 805, the processor performing process 800 rolls back to the last commit point. This effectively discards all speculative data accumulated to that point. After the rollback,  
15 in step 806, the processor enters into conservative execution mode. In step 807, execution is retried until the same exception point is reached.

If it is determined in step 803 that the execution mode is conservative, then in step 808, a cacheability character fix-up is executed appropriate for a cacheable-to-  
20 non-cacheable transition (e.g., Process 900; Figure 9). If it is determined in step 804 that the execution mode is conservative, then in step 809, a cacheability character fix-up is executed appropriate for a non-cacheable-to-cacheable transition (e.g., Process 1000; Figure 10). Process 800 can be complete upon completing steps 808 or 809.

25

Figure 9 is a flowchart of an exemplary computer implemented process 900 for handling a cacheability exception in a cacheable-to-non-cacheable transition,

according to one embodiment of the present invention. Process 900 begins with step 901, wherein old TLB entries having the affected physical address to be accessed are removed.

5           In step 902, a cacheability map (e.g., an authoritative table of cacheability characteristics indexed by page identifier) of the physical memory is updated. The cacheability map entry corresponding to the affected physical memory page is set to "non-cacheable."

10           In step 903, the cache is flushed of the affected physical page. In step 904, execution continues with entering the candidate TLB entry to the TLB and resuming the instruction that caused the original TLB miss, completing process 900.

            Figure 10 is a flowchart of an exemplary computer implemented process  
15   1000 for handling a cacheability exception in a non-cacheable-to-cacheable transition, according to one embodiment of the present invention. Process 1000 begins with step 1001, wherein all pending writes to the memory from the non-cacheable path (e.g., write combiners 16; Figure 1A) are drained. In one embodiment, to effectuate a desired weaker memory coherency model, step 1001  
20   can be optional.

            In step 1002, old TLB entries having the affected physical address to be accessed are removed. In step 1003, the cacheability map of the physical memory is updated. The cacheability map entry corresponding to the affected physical  
25   memory page is set to "cacheable."

In step 1004, execution continues with entering the candidate TLB entry to the TLB and resuming the instruction that caused the original TLB miss, completing process 1000.

5 In summary, a method and system for enforcing consistent per-physical page cacheability attributes, is disclosed. In one embodiment, the method for enforcing consistent per-physical page cacheability attributes maintains memory coherency within a processor addressing memory by comparing a desired cacheability attribute of a physical page address in a PTE against an authoritative table that indicates the  
10 current cacheability status. This comparison can be made at the time the PTE is inserted into a TLB. In one embodiment, when the comparison detects a mismatch between the desired cacheability attribute of the page and the page's current cacheability status, corrective action can be taken to transition the page into the desired cacheability state.

15 In one embodiment, the corrective action can be performed by either software or hardware. In one embodiment, the authoritative table resides within a private memory. The table is indexed with an identifier such as a page number, and provides an attributed current cacheability associated with that identifier. In one  
20 embodiment, a comparison process checks the page of memory whose cacheability attribute is being determined against the authoritative table and ascertains whether the page's desired cacheability attribute matches the current cacheability state indexed to that page or that a cacheability mismatch exists for that page. In one embodiment, where a cacheability mismatch is detected for a page,  
25 the cacheability attribute is modified. This modification in the page's cacheability attribute comprises a corrective action that functions to maintain cache coherency.

An embodiment of the present invention, a method and system for enforcing consistent per-physical page cacheability attributes, is thus described. While the present invention has been described in particular embodiments, it should be appreciated that the present invention should not be construed as limited by  
5 such embodiments, but rather construed according to the following claims and their equivalents.